

Stochastic method with low mode substitution for nucleon isovector matrix elements

Yi-Bo Yang¹, Andrei Alexandru², Terrence Draper¹, Ming Gong³, and Keh-Fei Liu¹



(χ QCD Collaboration)

¹*Department of Physics and Astronomy,
University of Kentucky, Lexington, KY 40506, USA*

²*Department of Physics,
The George Washington University,
Washington, DC 20052, USA*

³*Institute of High Energy Physics and Theoretical Physics Center for Science Facilities,
Chinese Academy of Sciences, Beijing 100049, China*

We introduce a stochastic method with low-mode substitution to evaluate the connected three-point functions. The isovector matrix elements of the nucleon for the axial-vector coupling g_A^3 , scalar couplings g_S^3 and the quark momentum fraction $\langle x \rangle_{u-d}$ are calculated with overlap fermion on 2+1 flavor domain-wall configurations on a $24^3 \times 64$ lattice at $m_\pi = 330$ MeV with lattice spacing $a = 0.114$ fm.

PACS numbers: 11.15.Ha, 12.38.Gc, 12.39.Mk

I. INTRODUCTION

The proton isovector-axial coupling g_A^3 and quark momentum fraction $\langle x \rangle_{u-d}$ are important benchmarks to check whether the systematic uncertainties of lattice QCD simulation, such as finite lattice spacing, finite volume, and chiral extrapolation, are under control, by a correct reproduction of the corresponding experimental results. Since the noisy disconnected insertion contribution to the isovector part of the nuclear matrix element is canceled between two degenerate flavors, the values are obtained solely from the connected insertion and thus are relatively cheaper to compute with high precision to be considered as benchmarks.

Most attempts have resulted in values $\sim 10\%$ below the experimental number for the axial-vector coupling [1–8], while a few claim that their results could be consistent with experiment [9–12]. For the quark momentum fraction $\langle x \rangle_{u-d}$, overestimation by $\sim 20 - 30\%$ is common in most of the calculations [3, 7, 14–16] except [8].

Recently, attention has been paid to lattice QCD calculation of the isovector scalar matrix element g_S^3 in the proton [2, 11, 17, 18] due to its role in constraining possible scalar interactions at the TeV scale [19].

In this work, we calculate the isovector matrix elements of the nucleon for the axial-vector and scalar couplings and the quark momentum fraction with the valence overlap fermion on 2 + 1 flavor domain-wall fermion (DWF) configurations [20]. Compared to simulations with other actions, the overlap fermion provides the best control of

the systematic errors since it is free of explicit chiral symmetry breaking and gives small $O(a^2)$ errors, whereas the numerical work is more costly.

In order to improve SNR, the 8-grid smeared Z_3 noise source with low-mode substitution (LMS) [23–27] has been applied to the hadron two point correlator on the $24^3 \times 64$ lattice [28] which improves the error of the nucleon mass of a point source by a factor of 7 and that of the 8-grid source without smearing by a factor of 2.5. In this work, we use a stochastic sandwich contraction method to remove the need of multiple inversions in the sink-sequential approach and use the current-sequential method for the low modes in the propagator between the current and the sink. This is an extension of the noise grid smeared source with LMS to the three point function. Such a many-to-all correlator with LMS is useful when the low-eigenmode contributions are important in the relevant time windows where the physical quantities are extracted.

The structure of the rest of the paper is organized as follows. The LMS technique with noise grid source for the non-zero momentum case of the two point correlation function is provided in Sec. II. Sec. III discusses the possibility of applying LMS on all the four quark propagators in the proton three-point function. The numerical details are provided in Sec. IV. In Sec. V, the results of isovector matrix elements of the nucleon for the axial-vector g_A^3 , the scalar coupling g_S^3 and the quark momentum fraction $\langle x \rangle_{u-d}$ are provided. A short summary and outlook are presented in Sec. VI.

II. LOW MODE SUBSTITUTION WITH MIXED MOMENTUM GRID SOURCE

Let's first consider the nucleon two-point function (2pt) with the interpolation field of the nucleon [29],

$$\begin{aligned}\chi_\alpha(x) &= \epsilon^{abc} \psi_\alpha^{(u)a}(x) \psi_\beta^{(u)b}(x) (\tilde{C})_{\beta\gamma} \psi_\gamma^{(d)c}(x) \\ \bar{\chi}_{\alpha'}(x) &= -\epsilon^{a'b'c'} \bar{\psi}_{\gamma'}^{(d)c'} (\tilde{C})_{\gamma'\beta'} \bar{\psi}_{\beta'}^{(u)b'}(x) \bar{\psi}_{\alpha'}^{(u)a'}(x),\end{aligned}\quad (1)$$

where $\tilde{C} \equiv C\gamma_5 = \gamma_2\gamma_4\gamma_5$ in the Pauli-Sakurai gamma-matrix convention, used throughout this work. There are two kinds of the Wick contractions so the 2pt of the nucleon can be constructed in terms of the point-to-point quark propagator S as

$$\begin{aligned}C(y, x; \Gamma; S^{(u)}, S^{(d)}, S^{(u)}) &= \langle \epsilon^{abc} \epsilon^{a'b'c'} \\ &\text{Tr} \left(\Gamma S^{(u)aa'}(y, x) \right) \text{Tr} \left(\underline{S}^{(d)bb'}(y, x) S^{(u)cc'}(y, x) \right) \rangle \\ &- \langle \epsilon^{abc} \epsilon^{a'b'c'} \text{Tr} \left(\Gamma S^{(u)ab'}(y, x) \underline{S}^{(d)ba'}(y, x) S^{(u)cc'}(y, x) \right) \rangle \\ &= \langle \epsilon^{abc} \epsilon^{a'b'c'} \\ &\text{Tr} \left(\Gamma S^{(u)aa'}(y, x) \right) \text{Tr} \left(\underline{S}^{(d)bb'}(y, x) S^{(u)cc'}(y, x) \right) \\ &+ \text{Tr} \left(\Gamma S^{(u)aa'}(y, x) \underline{S}^{(d)bb'}(y, x) S^{(u)cc'}(y, x) \right) \rangle\end{aligned}\quad (2)$$

where \underline{S} is defined as $(\tilde{C}S\tilde{C}^{-1})^T$ and Γ is the projection operator for the nucleon polarization.

The quark propagator S in the above equation is the inverse of the operator $(D_c + m)$ [30, 31], where D_c is defined in terms of the overlap operator and is chiral, i.e. $\{D_c, \gamma_5\} = 0$ [32]. The details will be discussed in Sec. IV. As in Ref. [21, 28], we use the low lying eigenvalues and eigenvectors of the overlap fermion, λ_i and $|i\rangle$, satisfying $D_c|i\rangle = \lambda_i|i\rangle$ to speed up the inversion and separate the propagator into its low-mode and high-mode parts,

$$\begin{aligned}S_L(y, x) &= \sum_{|\lambda_i| < \epsilon_c} \frac{1}{\lambda_i + m} |i\rangle_y \langle i|_x, \\ S_H(y, x) &= S(y, x) - \sum_{|\lambda_i| < \epsilon_c} \frac{1}{\lambda_i + m} |i\rangle_y \langle i|_x,\end{aligned}\quad (3)$$

with ϵ_c as the upper bound of the modulus of the eigenvalues.

The idea of using the Z_3 noise grid source is to tie the sources of the three quark propagators stochastically to each point (or a smeared point) on the grid so that one can have a multi-to-all correlator from one inversion. LMS for the quark propagator with Z_3 noise grid source (PropNG), be it point-grid (PG) [21] or smeared grid (SG) [28], has been used to improve the SNR for the nucleon correlator with significant success. This technique removes the gauge non-invariant contributions of the low-mode contributions (defined below) from the cases in which three propagators are from different source sites, and restores the benefit of using PropNG.

To construct the nucleon correlation function with LMS, PropNG $S_{NG}(y)$ should be split into its high-mode and low-mode pieces

$$\begin{aligned}S_{NG}(y) &= \sum_{x \in G} \theta(x) S(y, x) \\ &= S_{NG}^H(y) + \sum_{x \in G} \theta(x) S^L(y, x),\end{aligned}\quad (4)$$

with $S_{NG}^H(y) = \sum_{x \in G} \theta(x) S^H(y, x)$ and random Z_3 phases $\theta(x) \in \{1, e^{i\frac{2}{3}\pi}, e^{-i\frac{2}{3}\pi}\}$ for each point on a grid G .

As in Ref. [28], we can expand the nucleon correlation function $C(y, x; \Gamma; S_{NG}^{(u)}, S_{NG}^{(d)}, S_{NG}^{(u)})$ with the decomposition in Eq. (4) (ignoring the indices for the sink position y and the projection matrix Γ),

$$\begin{aligned}&C^{LMS}(S_{NG}, S_{NG}, S_{NG}) = \\ &= C(S_{NG}^H, S_{NG}^H, S_{NG}^H) + \sum_{x \in G} C(\theta(x) S^L(x), \theta(x) S^L(x), \theta(x) S^L(x)) \\ &+ C\left(\sum_{x \in G} \theta(x) S^L(x), S_{NG}^H, S_{NG}^H\right) + C\left(S_{NG}^H, \sum_{x \in G} \theta(x) S^L(x), S_{NG}^H\right) + C\left(S_{NG}^H, S_{NG}^H, \sum_{x \in G} \theta(x) S^L(x)\right) \\ &+ \sum_{x \in G} C(\theta(x) S^L(x), \theta(x) S^L(x), S_{NG}^H) + \sum_{x \in G} C(\theta(x) S^L(x), S_{NG}^H, \theta(x) S^L(x)) + \sum_{x \in G} C(S_{NG}^H, \theta(x) S^L(x), \theta(x) S^L(x)) \\ &= C_{ker}(S_{NG}^H, \sum_{x \in G} \theta(x) S^L(x)) + \sum_{x \in G} C_{ker}(\theta(x) S^L(x), S_{NG}^H)\end{aligned}\quad (5)$$

where

$$C_{ker}(S_1, S_2) = C(S_1, S_1, S_1) + C(S_2, S_1, S_1) \\ + C(S_1, S_2, S_1) + C(S_1, S_1, S_2). \quad (6)$$

The nucleon correlator with LMS here can be obtained from the one in Ref. [28] with just one more step. The low-mode propagator $\sum_{x \in G} \theta(x) S^L(y, x)$ is decomposed into several terms as in the very last term in the RHS of Eq. 5 to improve the SNR.

After the noise averaging, the nucleon correlation function with PropNG should be a stochastic estimate of the sum of nucleon correlators from each of the grid points, i.e.

$$\sum_{\vec{y}} C_{grid}(\vec{y}) = \sum_i \sum_{\vec{y}} C(\vec{y}, \vec{w}_i), \quad (7)$$

where the grid points \vec{w}_i are

$$\vec{w}_i \in (x_0 + m_x \Delta_x, y_0 + m_y \Delta_y, z_0 + m_z \Delta_z). \quad (8)$$

with $m_{x,y,z} = (0, 1, \dots, L_s/\Delta_{x,y,z})$ modulo the periodic boundary condition in the spatial directions. In this grid pattern, in addition to the zero momentum mode $(0,0,0)$, one can obtain non-zero momentum modes from the nucleon correlation function with PropNG. For example, for the PropNG with a regular ($\Delta_x = \Delta_y = \Delta_z = L_s/m$) grid, the momentum mode $p = (\pm n_1 m, \pm n_2 m, \pm n_3 m)$ ($n_{1,2,3}$ are integers) can be obtained. In this case, there is a phase factor which needs to be taken into account when the origin $w_0 = (x_0, y_0, z_0)$ is changed from configuration to configuration,

$$\sum_{\vec{y}} C_{grid}(\vec{y}) e^{-i \frac{2\pi}{L_s} \vec{y} \cdot p} = e^{-i \frac{2\pi}{L_s} w_0 \cdot p} \\ \sum_i \sum_{\vec{y}} C(\vec{y}, \vec{w}_i) e^{-i \frac{2\pi}{L_s} (\vec{y} - \vec{w}_i) \cdot p} e^{-i \frac{2\pi}{L_s} (\vec{w}_i - \vec{w}_0) \cdot (n_1, n_2, n_3)} \\ = e^{-i \frac{2\pi}{L_s} w_0 \cdot p} \sum_i \sum_{\vec{y}} C(\vec{y}, \vec{w}_i) e^{-i \frac{2\pi}{L_s} (\vec{y} - \vec{w}_i) \cdot p} \quad (9)$$

The exponential term in the second line with the exponent proportional to $\vec{w}_i - \vec{w}_0$ does not contribute, since all components of the latter are proportional to L_s/m and, as a result, the exponent is a multiple of 2π .

In order to obtain the other momentum modes, propagators with noise grid non-zero momentum source (PropNGM) are required. To cover a range of p^2 modes and minimize the effect of the rotation symmetry breaking due to the finite lattice spacing and volume, three kinds of PropNGM

$$S_{p_1}(y) = \sum_i \theta(\vec{w}_i) S(\vec{y}, \vec{w}_i) e^{i \frac{2\pi}{L_s} \vec{w}_i \cdot (1,0,0)}, \\ S_{p_2}(y) = \sum_i \theta(\vec{w}_i) S(\vec{y}, \vec{w}_i) e^{i \frac{2\pi}{L_s} \vec{w}_i \cdot (0,1,0)}, \\ S_{p_3}(y) = \sum_i \theta(\vec{w}_i) S(\vec{y}, \vec{w}_i) e^{i \frac{2\pi}{L_s} \vec{w}_i \cdot (0,0,1)} \quad (10)$$

and related inversions are required for the proton case. It is trivial to confirm that one can obtain a momentum mode like $(1,1,0)$ from the contraction $C(S_{p_1}, S_{p_2}, S_{NG})$, and $(1,1,1)$ from $C(S_{p_1}, S_{p_2}, S_{p_3})$.

To reduce the cost, we can combine these three kinds of PropNGM together as the mixed PropNGM,

$$S_p \equiv S_{p_1} + S_{p_2} + S_{p_3} \\ = \sum_i \theta(\vec{w}_i) S(\vec{y}, \vec{w}_i) (e^{i \frac{2\pi}{L_s} \vec{w}_i \cdot (1,0,0)} \\ + e^{i \frac{2\pi}{L_s} \vec{w}_i \cdot (0,1,0)} + e^{i \frac{2\pi}{L_s} \vec{w}_i \cdot (0,0,1)}), \quad (11)$$

with the origin of the grid $\vec{w}_0 = (x_0, y_0, z_0)$ to be selected randomly for each configuration.

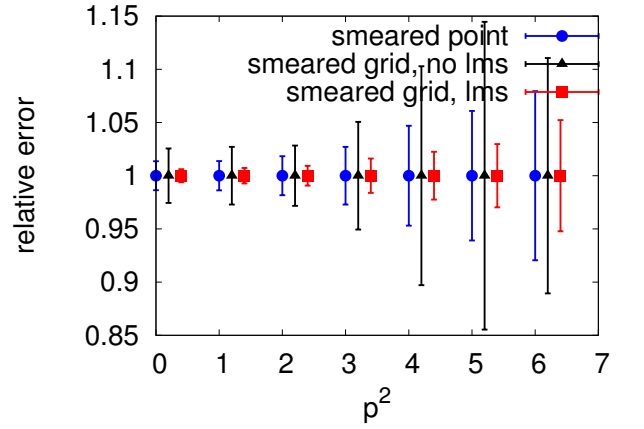


FIG. 1: The plot shows the relative error of 2pt as a function of the momentum squared p^2 at $t=8$ in lattice units. The data points of the smeared grid cases have been shifted a bit on the abscissa to make it easier to distinguish them. The SNR of the case with the noise smeared grid source (red squares) and LMS applied is better than the one with smeared point source (blue dots), while the one with the noise smeared grid source but no LMS (black triangles) is even worse than the one with smeared point source.

Fig. 1 shows the SNR of the proton effective mass at the unitary point where the pion mass due to the valence quark is the same as that from the sea, on the ensemble of which details will be addressed in Sec. IV. When LMS is applied, the SNR of the 2pt with the noise smeared grid source propagators (PropNG and mixed PropNGM, $\Delta_x = \Delta_y = \Delta_z = L_s/2$) is 2.3 times smaller than that of the smeared point source at $p^2 = 0$. This is a gain of 5.3 in statistics which is very good considering that the maximum possible gain is 8 for the ideal case where the independent nucleon propagators emerge from each of the 8 smeared grid points. On the other hand, if we don't use LMS, the SNR of 2pt with grid source is worse than the smeared point source, even though the latter has only 1/8 of the statistics of the former. This is understood as due to the fact that the Parisi-Lepage

estimate of the SNR for the nucleon is modified to

$$\frac{C_N(t, \vec{p}=0)}{\sigma_N(t)} \approx \sqrt{\frac{N}{V_3}} e^{-(m_N - 3/2 m_\pi)t}, \quad (12)$$

where N is the product of the number of noise and the number of gauge configurations and V_3 is the three-volume of the noise with its support on a time slice. In our case, $V_3 = 8$. It is this extra factor of $\frac{1}{\sqrt{V_3}}$ which makes the SRN of the 2pt from the noise smeared grid source without LMS worse than that of the smeared point source. When LMS is employed, the situation is reversed and one gains a statistical factor almost as large as the

number of the grid points. Thus, it is essential to have LMS when the noise grid source is used for the nucleon.

III. LMS OF THE CONNECTED THREE-POINT CORRELATOR

Generally, a nucleon three point function (3pt), from x to y , with a current $\bar{\psi}(x)^{(u)} \mathcal{O}(z) \psi(x)^{(u)}$ (with current operator \mathcal{O} such as γ_i , $\gamma_i D_j$, etc.) inserted at z , includes four kinds of Wick contractions,

$$\begin{aligned} C_3^u(y, x; \Gamma; \hat{S}^{(u)}, S^{(u)}, S^{(d)}, S^{(u)}) &= \langle \epsilon^{abc} \epsilon^{a'b'c'} \text{Tr} \left(\Gamma S^{(u)ad}(y, z) \mathcal{O}(z) S^{(u)da'}(z, x) \right) \text{Tr} \left(\underline{S}^{(d)bb'}(y, x) S^{(u)cc'}(y, x) \right) \rangle \\ &+ \langle \epsilon^{abc} \epsilon^{a'b'c'} \text{Tr} \left(\Gamma S^{(u)ad}(y, z) \mathcal{O}(z) S^{(u)da'}(z, x) \underline{S}^{(d)bb'}(y, x) S^{(u)cc'}(y, x) \right) \rangle \\ &+ \langle \epsilon^{abc} \epsilon^{a'b'c'} \text{Tr} \left(\Gamma S^{(u)aa'}(y, x) \right) \text{Tr} \left(\underline{S}^{(d)bb'}(y, x) S^{(u)cd}(y, z) \mathcal{O}(z) S^{(u)dc'}(z, x) \right) \rangle \\ &+ \langle \epsilon^{abc} \epsilon^{a'b'c'} \text{Tr} \left(\Gamma S^{(u)aa'}(y, x) \underline{S}^{(d)bb'}(y, x) S^{(u)cd}(y, z) \mathcal{O}(z) S^{(u)dc'}(z, x) \right) \rangle \end{aligned} \quad (13)$$

and can be expressed in terms of the 2pt correlation function $C(y, x; \Gamma; S^{(u)}, S^{(d)}, S^{(u)})$ defined in Eq. (2),

$$\begin{aligned} C_3^u(y, x; \Gamma; \hat{S}^{(u)}, S^{(u)}, S^{(d)}, S^{(u)}) \\ = C(y, x; \Gamma; \hat{S}^{(u)}, S^{(d)}, S^{(u)}) \\ + C(y, x; \Gamma; S^{(u)}, S^{(d)}, \hat{S}^{(u)}), \end{aligned} \quad (14)$$

where $\hat{S}(\mathcal{O}, z_0; y, x) \equiv \sum_z S(y, z) \mathcal{O}(z) S(z, x)$ is the current inserted propagator (PropCI). Similarly, the 3pt with a current of d quark can be expressed as

$$\begin{aligned} C_3^d(y, x; \Gamma; \hat{S}^{(d)}, S^{(u)}, S^{(d)}, S^{(u)}) \\ = C(y, x; \Gamma; S^{(u)}, \hat{S}^{(d)}, S^{(u)}). \end{aligned} \quad (15)$$

Fig. 2 shows PropCI as the product of the propagators in the shadowed region.

Supposing $S^{(u)} = S^{(d)} = S$, Eq. 14 can be rewritten into the contraction of PropCI \hat{S} and the remaining parts denoted as $X_{u,d}(\Gamma, S_1, S_2)$,

$$\begin{aligned} C_3^u(\Gamma; \hat{S}, S, S, S) &= \langle \text{Tr}(\hat{S} X_u(\Gamma, S, S)) \rangle, \\ C_3^d(\Gamma; \hat{S}, S, S, S) &= \langle \text{Tr}(\hat{S} X_d(\Gamma, S, S)) \rangle, \end{aligned} \quad (16)$$

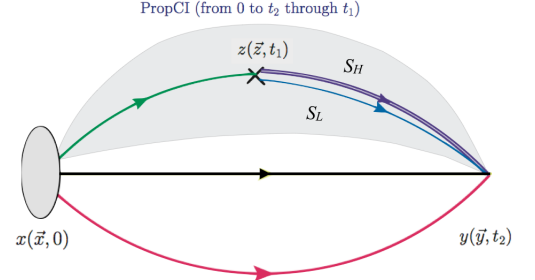


FIG. 2: The quark diagram of the proton correlation function with the connected insertion, from x to y , with an insertion at z . The product of the propagators in the shadowed region is the current inserted propagator, \hat{S} . The propagator from the current z to the sink y is decomposed into its low- and high-mode contributions (S_L and S_H respectively) for further SNR/cost improvement from the advanced technique in the latter discussion. See Sec. III B for more details.

with

$$\begin{aligned}
X_u^{aa'}(\Gamma, S_1, S_2) &= \epsilon^{abc} \epsilon^{a'b'c'} \\
&\quad (\Gamma \text{Tr}[\underline{S}_2^{bb'} S_1^{cc'}] + \underline{S}_2^{bb'} S_1^{cc'} \Gamma \\
&\quad + \text{Tr}[\Gamma S_1^{cc'}] \underline{S}_2^{bb'} + \Gamma S_1^{cc'} \underline{S}_2^{bb'}), \\
X_d^{bb'}(\Gamma, S_1, S_2) &= \epsilon^{abc} \epsilon^{a'b'c'} \\
&\quad (\text{Tr}[\Gamma S_1^{aa'}] \tilde{C}^{-1}(S_2^{cc'})^T \tilde{C} \\
&\quad + \tilde{C}^{-1}(S_1^{aa'} \Gamma S_2^{cc'})^T \tilde{C}) \quad (17)
\end{aligned}$$

Based on the above definition, a typical 3pt correlation function for a point source on the $t = 0$ time slice, when summed over the spatial indices of y and z becomes

$$\begin{aligned}
C_3(t_2, t_1) &= \sum_{\vec{y}} \langle \text{Tr}[\hat{S}(\mathcal{O}, t_1; \vec{y}, t_2, \vec{0}, 0) \\
&\quad X_{u,d}(\vec{y}, t_2, \vec{0}, 0; \Gamma, S, S)] \rangle. \quad (18)
\end{aligned}$$

A. Sink-sequential method and Stochastic sandwich method

The typical problem of the connected 3pt is calculating the propagator from the current to the sink $S(\vec{y}, t_2, \vec{z}, t_1)$. On the surface, it is an all-to-all propagator which would be beyond the ability of the standard lattice inversion operation.

However, when the sink time t_2 is fixed, the sequential source method [33, 34] could be used, with $\gamma_5 X_{u,d}^\dagger(\vec{y}, t_2, \vec{0}, 0) \gamma_5$ as the source of the matrix inversion, to construct

$$\begin{aligned}
S_{seq}(X_{u,d}; \vec{z}, t_1, t_2, \vec{0}, 0) &= \sum_{\vec{y}} S(\vec{z}, t_1, \vec{y}, t_2) \\
&\quad \gamma_5 X_{u,d}^\dagger(\vec{y}, t_2, \vec{0}, 0) \gamma_5. \quad (19)
\end{aligned}$$

Then, one can contract S_{seq} with the standard quark propagator from $t = 0$ to t_1 to construct the 3pt correlator,

$$\begin{aligned}
C_3(t_2, t_1, \mathcal{O}) &= \sum_{\vec{z}, i} \text{Tr}[\gamma_5 S_{seq}^\dagger(X_{u,d}, \vec{z}, t_1, t_2, \vec{0}, 0) \gamma_5 \\
&\quad \mathcal{O}(\vec{z}, t_1) S(\vec{z}, t_1, \vec{0}, 0)], \quad (20)
\end{aligned}$$

taking the advantage of the relation $\gamma_5 S(z, y)^\dagger \gamma_5 = S(y, z)$.

The disadvantage of the sequential method is that it has to calculate the sink-sequential propagator repeatedly when X is changed for any reason, such as for: different momentum, different quark flavor or mass, or different polarization projection of the baryon. This is expensive when many momenta are needed.

The number of inversions required in the sink-sequential method is $2 \times 4 \times N_p$ where the 2 is for the u and d flavors in the nucleon, 4 is for the polarization, and N_p is the number of momentum projections. When

many N_p are required for nucleon form factors with momentum transfer (hundreds are needed for $|\vec{p}| \leq 3$ with high statistics), the cost can be staggering.

A stochastic method [37–39] (referred to as the stochastic sandwich method (SSM) in this work) is introduced to reduce the cost of the sequential method when many sequential inversions are required. It entails inserting a noise estimate of the delta function $\delta(\vec{y}_1, \vec{y}_2)$ at $t = t_2$,

$$\begin{aligned}
\frac{1}{N_{noi}} \sum_{i=1}^{N_{noi}} \sum_{\vec{y}_1, \vec{y}_2, \vec{z}} \text{Tr}[\theta_{\vec{y}_1}^{(i)} S(\vec{y}_1, t_2, \vec{z}, t_1) \mathcal{O}(\vec{z}, t_1) S(\vec{z}, t_1, \vec{0}, 0) \\
X(\vec{0}, 0, \vec{y}_2, t_2) \theta_{\vec{y}_2}^{(i)\dagger}] \xrightarrow{N_{noi} \rightarrow \infty} C_3(t_2, t_1, \mathcal{O}), \quad (21)
\end{aligned}$$

where N_{noi} is the number of the noises and the noise θ satisfies

$$\frac{1}{N_{noi}} \sum_{i=1}^{N_{noi}} \theta_{\vec{y}_1}^{(i)} \theta_{\vec{y}_2}^{(i)\dagger} \xrightarrow{N_{noi} \rightarrow \infty} \delta_{\vec{y}_1, \vec{y}_2}. \quad (22)$$

In other words, it uses the noise estimate of the all-to-all propagator,

$$S(\vec{y}_1, t_2, \vec{z}, t_1) \cong \sum_i \theta_{\vec{y}_1}^{(i)} \gamma_5 (S_{noi}^{(i)}(\vec{z}, t_1, t_2))^\dagger \gamma_5 \quad (23)$$

with

$$S_{noi}^{(i)}(\vec{z}, t_1, t_2) = \sum_{\vec{y}_1} S(\vec{z}, t_1, \vec{y}_1, t_2) \theta_{\vec{y}_1}^{(i)\dagger}, \quad (24)$$

instead of the original $S(\vec{y}, t_2, \vec{z}, t_1)$, to avoid the expensive calculation to construct the sink-sequential propagator with inversion of $2 \times 4 \times N_p$ sources.

B. Stochastic sandwich method (SSM) with LMS

SSM avoids the cost of the repeated inversion for many different sequential sources, but it still requires multiple inversions for several noises, before the SNR can reach its upper limit – that of the sequential method. In this work, the basic idea is to improve the SNR of the 3pt correlator of SSM using the low lying eigenvectors of D_c to construct the long distance part of the all-to-all $S(\vec{y}, t_2, \vec{z}, t_1)$ (S_L in Fig. 2, the single line from the current to the sink), and using the noise many-to-all propagator to estimate the remaining high frequency part of $S(\vec{y}, t_2, \vec{z}, t_1)$ (S_H in Fig. 2, the double line from the current to the sink). Thus, the propagator with LMS is written as

$$\begin{aligned}
S^{LMS}(\vec{y}_1, t_2, \vec{z}, t_1) &= \sum_i \theta_{\vec{y}_1}^{(i)} \gamma_5 (S_{noi}^{(i),H}(\vec{z}, t_1, t_2))^\dagger \gamma_5 \\
&\quad + \sum_i \frac{1}{\lambda_i + m} v_i(\vec{y}, t_2) v_i^\dagger(\vec{z}, t_1). \quad (25)
\end{aligned}$$

where λ_i and v_i are the low-lying eigenvalues and the corresponding eigenvectors of D_c . In other words, it is

a technique to apply LMS to the sequential propagator $S_{seq}(X_{u,d}; \vec{z}, t_1, t_2, \vec{0}, 0)$ (LMS_S). It is expected to reduce the number of the noise propagators needed to reach the upper limit of SNR.

When LMS_S in Eq. (25) is applied to the PropCI in Eq. (14), \hat{S} comes from $t = 0$ to $t = t_2$ through $t = t_1$

$$\begin{aligned} \hat{S}^{LMS_S}(\mathcal{O}, t_1; \vec{y}, t_2, t_1, \vec{0}, 0) &= \\ &= \sum_{\vec{z}} S^{LMS_S}(\vec{y}_1, t_2, \vec{z}, t_1) \mathcal{O}(\vec{z}, t_1) S(\vec{z}, t_1, \vec{0}, 0) \\ &= \sum_{\vec{z}} \left(\sum_i \frac{1}{\lambda_i + m} v_i(\vec{y}, t_2) v_i^\dagger(\vec{z}, t_1) \right. \\ &\quad \left. + \theta^{(i)}(\vec{y}, t_2) \sum_{\vec{z}, i} \gamma_5 (S_{noi}^{(i), H}(\vec{z}, t_1, t_2))^\dagger \gamma_5 \right) \\ &\quad \mathcal{O}(\vec{z}, t_1) S(\vec{z}, t_1, \vec{0}, 0), \end{aligned} \quad (26)$$

as shown in the shadowed area in Fig. 2.

Then one can construct 3pt with LMS by constructing the standard 2pt repeatedly (the projection matrix Γ is suppressed for clarity),

$$\begin{aligned} C_3^{LMS, u}(\hat{S}, S) &= C_{ker}^3(\hat{S}^H, \hat{S}^L, S_{NG}^H, S_{NG}^L, S_{NG}^H, S_{NG}^L) + \sum_{x \in G} C_{ker}^3(\hat{S}^{L(x)}, \hat{S}^H, \theta(x) S^L, S_{NG}^H, \theta(x) S^L, S_{NG}^H) + \\ &\quad C_{ker}^3(S_{NG}^H, S_{NG}^L, S_{NG}^H, S_{NG}^L, \hat{S}^H, \hat{S}^L) + \sum_{x \in G} C_{ker}^3(\theta(x) S^L, S_{NG}^H, \theta(x) S^L, S_{NG}^H, \hat{S}^{L(x)}, \hat{S}^H) \\ C_3^{LMS, d}(\hat{S}, S) &= C_{ker}^3(S_{NG}^H, S_{NG}^L, \hat{S}^H, \hat{S}^L, S_{NG}^H, S_{NG}^L) + \sum_{x \in G} C_{ker}^3(\theta(x) S^L, S_{NG}^H, \hat{S}^{L(x)}, \hat{S}^H, \theta(x) S^L, S_{NG}^H) \end{aligned} \quad (27)$$

where

$$\begin{aligned} S_{NG}^L &= \sum_{x \in G} \theta(x) S^L(x), \text{ and} \\ C_{ker}^3(X_1, X_2, Y_1, Y_2, Z_1, Z_2) &= C(X_1, Y_1, Z_1) \\ &\quad + C(X_2, Y_1, Z_1) + C(X_1, Y_2, Z_1) + C(X_1, Y_1, Z_2) \end{aligned} \quad (28)$$

and \hat{S}^H and $\hat{S}^{L(x)}$ are the high- and low-mode parts of \hat{S}^{LMS_S} in Eq. (26).

This is the stochastic sandwich method with LMS which uses the low eigenmodes for the propagator from the current to the sink in PropCI, \hat{S}^{LMS_S} with current insertion and the high modes for the same which originates from the sink time slice. The construction of the PropCI with low modes needs to be done for each current and momentum transfer and t_2 (if desired). In contrast, the current-sequential method will need to do an inversion for each current, momentum transfer, and t_1 separately.

To account for the amount of numerical work for different approaches to the 3pt CI correlators, we note the traditional sink-sequential method entails $2 \times 4 \times N_p$ inversions at a fixed sink time slice t_2 , where the 2 and 4 refer to the separate sources X in Eq. (17) labeled with u and d flavors and polarization directions (unpolarized and polarization in 3 spatial directions). N_p is the number of sink momenta for the nucleon. For SSM without LMS,

there are N_{noi} inversions of the N_{noi} noise vectors at the sink time t_2 . How many N_{noi} is needed for acceptable SNR depends on the observable. For the SSM with LMS, besides the noise propagator S_{noi}^H with N_{noi}^H inversion, there is an overhead for the low-mode portion of PropCI (\hat{S}^{LMS_S} in Eq. (26)). It includes N times the low-mode contributions from N smeared grid source plus one high-mode contribution for the propagator from the source to the current (S_{NG}^H). Each needs to be folded with the current for different momentum transfer \vec{q} . Therefore the overhead is $\epsilon \times (N + 1) \times N_{cu} \times N_q$ where N_{cu}/N_q is the number of currents/momentum transfer, and ϵ is the fraction of inversion time for constructing the low-mode portion of \hat{S}^{LMS_S} for each current and momentum transfer. We list the cost for the sink and current parts of the 3pt function in units of quark inversion in Table I for future reference. To evaluate the efficacy among the three methods, one needs to compare costs in the table to reach the same precision for a given observable. For the case of SSM with LMS, there is an additional gain from the noise grid source with LMS as discussed in Sec. II which needs to be taken into account.

TABLE I: The cost for the sink and current parts of the 3pt function in units of quark inversion is listed for the sink-sequential method (Sequential), stochastic sandwich method (SSM), and SSM with LMS. N_p is the number of sink nucleon momenta, N_{noi} is the number of noise in SSM. N_{noi}^H is the number of noise in SSM with LMS, and N_{cu}/N_q is the number of currents/momentum transfer in the construction of the low-mode part of PropCI. ϵ is the fraction of inversion time for constructing the low-mode portion of PropCI for each current and momentum transfer and N_p momenta (~ 0.02 on the ensemble used in this work).

Sequential	SSM	SSM+LMS _S
$8N_p$	N_{noi}	$N_{noi}^H + \epsilon(N+1)N_{cu}N_q$

IV. NUMERICAL DETAILS

In this work, we use the valence overlap fermion on $2+1$ flavor domain-wall fermion (DWF) configurations [20] to carry out the calculation [21].

The lattice we use has a size $24^3 \times 64$ with lattice spacing $a^{-1} = 1.75(4)$ GeV set by r_0 at the chiral and continuum limits [42]. The light sea u/d quark mass $m_l a = 0.005$ corresponds to $m_\pi \sim 330$ MeV. We have calculated the isovector matrix elements of the nucleon for the axial-vector and scalar couplings and the quark momentum fraction at 6 valence quark mass parameters which correspond to the renormalized masses $m_q^R \equiv m_q^{\overline{\text{MS}}}(2\text{GeV})$ ranging from 13 to 32 MeV after the non-perturbative renormalization procedure in Ref. [41]. They correspond to the pion mass in the range of 250-400 MeV. In order to enhance the signal-to-noise ratio in the calculation of three-point functions, we use two smeared noise 12-12-12 grid sources at $t_i = 0$ and 32 (one is PropNG and the one is PropNGM) [28] and two noise 2-2-2 grid point sources at positions t_f which are 8, 10, and 12 time-slices away from the sources on 203 configurations.

The effective overlap operator D_c is chiral, i.e. $\{D_c, \gamma_5\} = 0$ [32], and is expressed in terms of the overlap operator D_{ov} as

$$D_c = \frac{\rho D_{ov}}{1 - D_{ov}/2} \text{ with } D_{ov} = 1 + \gamma_5 \epsilon(\gamma_5 D_w(\rho)), \quad (29)$$

where ϵ is the matrix sign function and D_w is the Wilson Dirac operator with a negative mass characterized by the parameter $\rho = 4 - 1/2\kappa$ for $\kappa_c < \kappa < 0.25$. We set $\kappa=0.2$ which corresponds to $\rho = 1.5$.

Compared to the earlier implementation of the overlap operator [21], the current implementation further improves the performance of data exchange on different nodes of the cluster and uses the polynomial approximation for the overlap operator instead of the rational approximation, and has achieved better scaling and further speed up of the calculation by a factor of two on average [22].

The number of D_c 's low mode eigenvectors used for the deflation of the overlap operator inversion and LMS, on this $24^3 \times 64$ lattice, is 200 pairs plus the zero modes, and the upper bound of the absolute value of the eigenvalues is 0.154 which is over two times larger than the dimensionless strange quark mass.

We check the efficacy of the sequential low-mode substitution (LMS_S) in the PropCI by examining the 3pt functions for the isovector axial and scalar currents. We plot the ratio of 3pt-to-2pt correlators as a function of the current insertion time t_1 in Fig. 3 where the sink time t_2 is 10. The blue dots and black triangles show the contributions where the current-to-sink part of PropCI is from the low modes and the noise-estimated high modes respectively. Notice that the contribution from the low modes is much larger than that of the high modes when the current time slice is farther away from the sink (i.e. closer to the source with small t_1) for both the axial and scalar cases, which reflects the fact that the low modes dominate the long-distance behavior of the PropCI between t_1 to t_2 . When the current is closer to the sink with larger t_1 , we see that the high modes dominate for the axial case which shows that the high modes are important and dominate the short distance behavior of the propagator. However, the high-mode contribution is still small for the scalar current case when t_1 is close to the sink which shows that the high-mode contribution is small for the 3pt function for the scalar current.

The red squares are the sum of the low- and high-mode contributions from the present hybrid scheme. We have also calculated the 3pt function without LMS_S for the PropCI, but instead use only the noise propagator as the full propagator from t_1 to t_2 . These are shown as the green triangles in Fig. 3. These correspond to the stochastic method introduced by the QCDSF Collaboration [37, 38] and the Cyprus group [39]. Since our LMS_S replaces the long distance part of the current-to-sink part of PropCI with an exact all-to-all one, the larger its contribution the larger the improvement. As in Fig. 3, the blue dots contribute over 80% in the g_S^u case and so the improvement of LMS_S is larger than in the g_A^u case. The error bars of SSM at the time slices $t_1 = 2 - 6$ turn out to be a factor ~ 2 for g_A^u (~ 4 for g_S^u) larger than that of using LMS_S in the present approach.

The fact that the error of g_A^3/g_S^3 in our approach is smaller than that of SSM with 2 noises by a factor of $\sim 2/4$ shows that it would take 8/32 noise inversions for SSM to have the same error as the present method with LMS. To compare the cost of SSM + LMS, we should take its overhead into account. On the present lattice, the percentage of inversion time for low-mode construction is $\epsilon = 0.02$. Therefore, the overhead $\epsilon(N+1)N_{cu}N_q = 0.72$ for $N = 8$ (smeared grid source), $N_{cu} = 4$ to account for the scalar current and A_i for 3 spatial directions and $N_q = 1$. Together with $N_{noi}^H = 2$, the cost is 2.72 inversions. This means that, to reach the same error, it would take SSM 2.9 and 11.8 times more inversions than SSM with LMS for g_A^3 and g_S^3 respectively. Furthermore, the

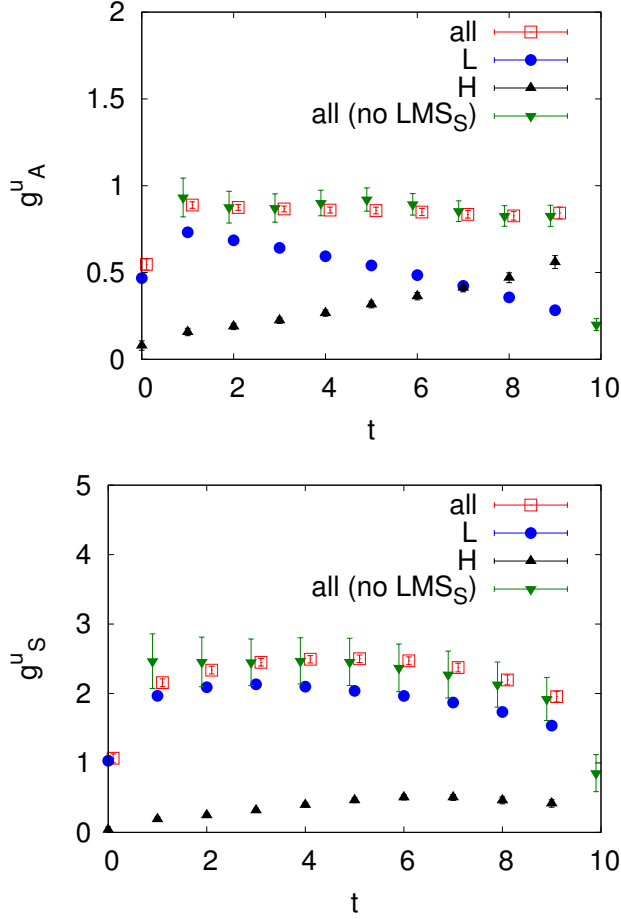


FIG. 3: The 3pt-to-2pt ratio with LMS_S (red squares) vs. the one without it (green inverted triangles). The source/sink is located at 0/10, and the current dependence for the matrix element with the current-to-sink part of PropCV including just the low- or high-mode parts are plotted as the blue dots or black triangles respectively. The upper panel is for the axial-vector current case and the lower panel is for the scalar case. Notice that the contribution of the low-mode part is larger when the current time slice is farther away from the sink.

smeared grid source with LMS has improved the statistics by a factor of 5.3 for $N = 8$ for the 2pt function. This additional factor of improvement is also expected for the 3pt function.

To compare with the sink-sequential method, we assume that our results have reached the SNR of that of the sink-sequential method. This is consistent with the fact that in the range $t_1 = 2-6$ where the observables are fitted, the PropCI are dominated by the low-mode contributions, particularly for g_S^3 . In this case, the cost of sink-sequential takes 16 inversions. Here, we have taken $N_p = 2$ to include the $\langle x \rangle_{u-d}$ calculation in addition to g_A^3 and g_S^3 . For the overhead in SSM + LMS, the number of currents needed is $N_{cu} = 6$ for these three quantities

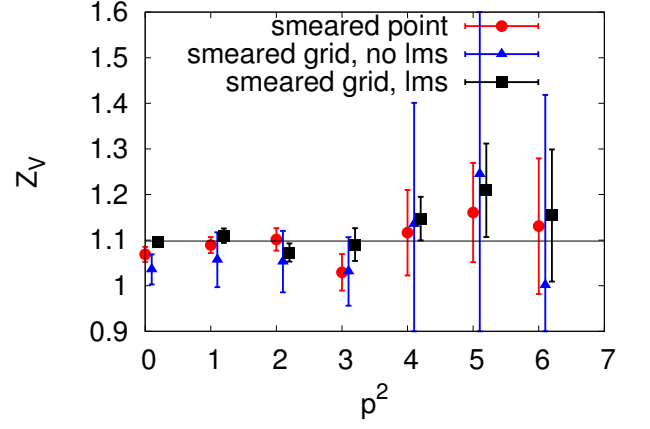


FIG. 4: The vector renormalization constant in the rest/moving frame at the unitary point, as a function of the momentum squared p^2 in lattice unit. The $p^2 = 0, 4$ involve PropNG only and the other cases involve PropNGM also. The former case gains more from LMS (black squares vs. red dots). The results obtained using the 8-point grid source without LMS applied are very noisy (blue triangles).

and the overhead is $\epsilon(N + 1)N_{cu}N_q = 1.08$. Therefore, besides the improvement from use of the grid source, the present method would be $16/3.08 = 5.2$ times more efficient than the sink-sequential method for the calculation of the three quantities. Note that the cost of the sink-sequential method has additional factors that need to be taken into account, such as N_{mass} for different masses, and also N when the necessary LMS is applied on the source of the sink-sequential propagator (as in Eq. 17), so SSM is much cheaper than the sink-sequential method.

When the physical volume is increased, while keeping the lattice spacing unchanged, and with a noise vector covering the entire spatial volume of the sink time slice, we expect that the region essentially contributing to 3pt will not change, while the remaining region contributes only to the noise. Such a simple argument hints that the noise required to reach the same SNR is proportional to volume and we have confirmed it explicitly on the $48^3 \times 96$ lattice with similar lattice spacing [40]. At the same time, the number of low modes will be proportional to volume if we want to reach the same upper bound of the eigenvalues, so the SSM with LMS will not lose its efficiency as compared to SSM without LMS, when the volume is larger. But, since the number of inversions is fixed in the standard sequential method, the SSM with and without LMS will lose their comparative efficiencies when the volume is very large.

Another issue we need to check is the effect of LMS in the 3pt case. For the 3pt function, we check, for example, the vector charge renormalization constant from the forward matrix element at the unitary point for several nucleon momenta. For $p^2 = 0$ and 4, only the propagator PropNG is involved, while the other cases involve Prop-

NGM also. In the former cases, we find that the smeared grid source with LMS improves the SNR by a factor of 2.0 compared to that with a smeared point source without LMS, slightly smaller than what we found with the 2pt function as discussed in Sec. II; whereas, the gain is only 1.4 for the other p^2 where the PropNGM is involved. We shall look into the possibility of improving the SNR further when PropNGM is involved.

V. RESULTS

A standard 3pt/2pt ratio in the forward matrix element case is

$$\begin{aligned}
 R(t_2, t_1, 0) &= C_3(t_2, t_1, 0)/C(t_2, 0) \\
 &= \frac{\sum_{i,j} Z_f^{(i)} Z_i^{(j)} e^{-E^{(i)}(t_2-t_1)-E^{(j)}t_1} \langle \chi_f^{(i)} | J | \chi_i^{(j)} \rangle}{\sum_k Z_f^{(k)} Z_i^{(k)} e^{-E^{(k)}t_2}} \\
 &\xrightarrow{t_2 \gg 0} \langle \chi_f^{(0)} | J | \chi_i^{(0)} \rangle \\
 &\quad + \frac{Z_f^{(1)}}{Z_f^{(0)}} \langle \chi_f^{(1)} | J | \chi_i^{(0)} \rangle e^{-\Delta E(t_2-t_1)} \\
 &\quad + \frac{Z_i^{(1)}}{Z_i^{(0)}} \langle \chi_f^{(0)} | J | \chi_i^{(1)} \rangle e^{-\Delta E t_1} \\
 &\quad + \frac{Z_f^{(1)} Z_i^{(1)}}{Z_f^{(0)} Z_i^{(0)}} (\langle \chi_f^{(1)} | J | \chi_i^{(1)} \rangle - \langle \chi_f^{(0)} | J | \chi_i^{(0)} \rangle) e^{-\Delta E t_2} \\
 &\quad + \dots,
 \end{aligned} \tag{30}$$

where $E^{(i)}$ and $Z^{(i)}$ are the energy and the overlap of the interpolation field of the i th state and $\Delta E = E^{(1)} - E^{(0)}$. For $t_2 \gg t_1 \gg 0$, the contributions from all the terms in the right hand of Eq. (30) except the first term vanish, and then one can use Eq. (30) to obtain the matrix element.

When t_2 is fixed, one may fit the first term and the combined second and third terms around $t_1 = t_2/2$ to include the effect of the ground state to first excited state transition in the right hand side of Eq. (30) which is t_1 dependent. But since the fourth term in the right hand side of Eq. (30), which is the difference of the matrix element in the ground state and the first excited state, is independent of t_1 just like the first term, one will not be able to disentangle them and, as a result, a systematic error may be induced by its contribution which is suppressed by $e^{-(E^{(1)}-E^{(0)})t_2}$. To get a feeling for the size of the correction, let us suppose that the first excited state matrix element $\langle \chi_f^{(1)} | J | \chi_i^{(1)} \rangle$ is 30% different from the ground state matrix element $\langle \chi_f^{(0)} | J | \chi_i^{(0)} \rangle$, and the mass difference of the first excited state and the ground state is about 500 MeV. Then the correction from such an effect with $t_2=8, 10$ and 12 (with the nucleon source set at $t_0 = 0$) is about 3%, 2% and 1% respectively. To assess this error, we shall calculate the 3pt function at

three values of t_2 so that we can fit all four terms in Eq. (30).

In order to check the t_2 dependence of the plateau, three sets of propagators with two noise-grid point sources each at positions $t_2 = 8, 10$ and 12 time-slices away from the nucleon source are generated, and all the t_1 dependence of these three cases are plotted together for comparison in Fig. 5 for the vector current case. The sink-source separation dependence seems to be mild here, but in general the minimum separation required by other quantities can be different.

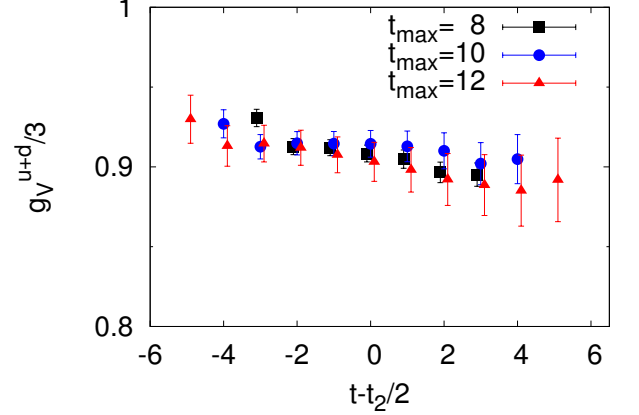


FIG. 5: The nucleon sink-source separation dependence of the matrix element with the vector charge for $u + d$ in the connected insertion. Obviously, the larger t_2 , the worse the signal. The data points marked with the black squares ($t_2=8$), the blue dots ($t_2=10$) and the red triangles ($t_2=12$) are consistent.

To check the separation effect quantitatively, we applied three kinds of fits to deduce the results:

The first method is to fit the ratio as a function of t_1 and t_2 ,

$$\begin{aligned}
 R^{fit}(t_2, t_1) &= C_0 + C_1 e^{-\Delta m(t_2-t_1)} \\
 &\quad + C_2 e^{-\Delta m t_1} + C_3 e^{-\Delta m t_2}
 \end{aligned} \tag{31}$$

with $C_{0,1,2,3}$ and Δm as free parameters. C_0 is the ground state matrix element we want. Since the t_1 dependence of $R(t_2, t_1)$ is mild in some of the quantities like g_V and g_A , we take Δm as a common parameter for all the quantities. This is what we mark as “2-state” in the following discussion.

In this work, we use the smeared source and the point sink, so the excited-state contaminations are different in the smaller and larger t ends. If the smeared source makes the contaminations in the smaller t end small, or has a different sign compared to that in the larger t end, the position of the plateau will be harder to determine, as in the case of g_V^{u+d} (Fig. 5) and g_A^3 (Fig. 7). Applying the “2-state” fit on such a quantity is not stable and provides large uncertainties (and/or large $\chi^2/d.o.f.$)

on the results. In this work, we constrain the mass difference Δm to be the same for the different matrix elements with the same quark mass value, and apply a correlated joint 2-state fit. To suppress the contamination from the excited state, we excluded the data points with $t_1 = 0, t_2 - 1$ and t_2 . One more data point at the larger t end is excluded since the excited-state contamination is larger there. Despite this, the fit is still not very good. Taking the unitary point as an example, the $\chi^2/d.o.f.$ with ~ 70 degrees of freedom is 1.45, the corresponding p-value is just 0.008. In addition, this method requires a joint fit with several quantities and is not suitable for the analysis of a single quantity without the information of the other quantities.

The second method is the sum method [43, 44] which is used in the disconnected insertion case, wherein a sum is taken over all the 3pt/2pt ratios in Eq. (30) with different t_1 ,

$$\begin{aligned}
 SR(t_2, t_1, 0) &= \sum_{0 < t_1 < t_2} R(t_2, t_1, 0) \\
 &= (t_2 - 1) \langle \chi_f^{(0)} | J | \chi_i^{(0)} \rangle \\
 &+ \frac{e^{-\Delta m}}{1 - e^{-\Delta m}} \left(\frac{Z_f^{(1)}}{Z_f^{(0)}} \langle \chi_f^{(1)} | J | \chi_i^{(0)} \rangle + \frac{Z_i^{(1)}}{Z_i^{(0)}} \langle \chi_f^{(0)} | J | \chi_i^{(1)} \rangle \right) \\
 &+ (t_2 - 1) \frac{Z_f^{(1)} Z_i^{(1)}}{Z_f^{(0)} Z_i^{(0)}} (\langle \chi_f^{(1)} | J | \chi_i^{(1)} \rangle - \langle \chi_f^{(0)} | J | \chi_i^{(0)} \rangle) \\
 &e^{-\Delta m t_2} + \dots
 \end{aligned} \tag{32}$$

When t_2 is large, we can use the linear function of t_2 (ignoring the $e^{-\Delta m t_2}$ correction)

$$SR^{fit}(t_2, t_1) = t_2 C_0 + C'_1 \tag{33}$$

to fit our summed ratio with 3 different separations, and obtain the slope as the ground state matrix element. This method will be marked as “sum” in the following discussion.

We found that the “sum” fit can obtain a $\chi^2/d.o.f.$ smaller than one, for all the quantities. But this fit just has one degree of freedom. Ignoring the $e^{-\Delta m t_2}$ correction can induce an uncontrolled systematic error.

The third method is to combine the first two methods, by fitting both the ratios and their sum together (denoted as “mixed”),

$$\begin{aligned}
 R^{fit}(t_2, t_1) &= C_0 + C_1 e^{-\Delta m(t_2 - t_1)} \\
 &+ C_2 e^{-\Delta m t_1} + C_3 e^{-\Delta m t_2},
 \end{aligned} \tag{34}$$

$$\begin{aligned}
 SR^{fit}(t_2, t_1) &= t_2 C_0 + \frac{e^{-\Delta m}}{1 - e^{-\Delta m}} (C_1 + C_2) \\
 &+ (t_2 - 1) C_3 e^{-\Delta m t_2} + C_4,
 \end{aligned} \tag{35}$$

where $C_{0,1,2,3}$ and Δm are the same as that in the “2-state” fit, and C_4 is for the constant contribution from the transition between higher excited states and the ground state.

The “2-state” fit makes fully use of the ratios, while it is unstable when the position of the plateau is hard to determine (such as for g_A^3). The “sum” fit provides a stable estimate of the ground state matrix element, but it suffers from the systematic error from ignoring the $e^{-\Delta m t_2}$ correction. By combining them together, we can obtain a stable fit of all the quantities discussed in this work independently, and don’t have to use a joint fit with several quantities. The $\chi^2/d.o.f.$ of different quantities and quark masses vary between 1.0 and 1.5 with 18 degrees of freedom, corresponding to p-values in the range of [0.08-0.46]. The value of Δm we obtained at the unitary point has a strong dependence on the quantity and varies from 400 MeV to 1 GeV.

The values for the renormalized isovector axial vector coupling g_A^3 , scalar coupling g_S^3 and quark momentum fraction $\langle x \rangle_{u-d}$ from the three methods at the unitary point are listed in Table II.

TABLE II: Isovector axial-vector coupling g_A^3 , scalar coupling g_S^3 and quark momentum fraction $\langle x \rangle_{u-d}^E$ at the unitary point from three fitting methods. See the following three subsections for the details.

	2-state	sum	mixed
g_A^3	1.189(20)	1.157(18)	1.166(19)
g_S^3	0.61(6)	0.78(6)	0.74(4)
$\langle x \rangle_{u-d}^E$	0.209(12)	0.190(13)	0.193(19)

A. Vector and Axial vector case

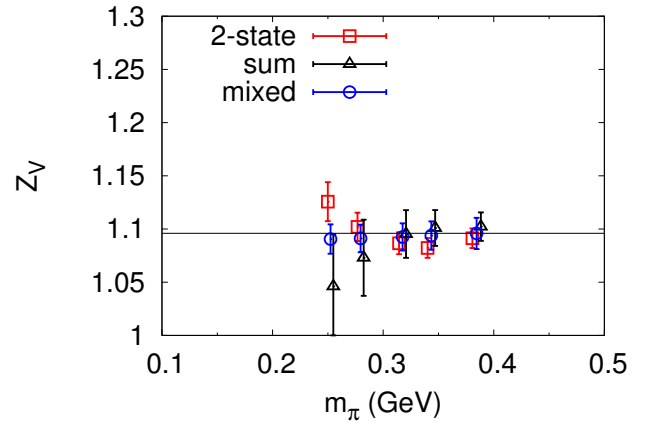


FIG. 6: The vector renormalization factor from the charge vs the pion mass, from three kinds of fitting methods: 2-state fit (red squares), summed slope (black triangles), and the mixed fit which combines those two methods (blue dots). The results from these different methods are consistent while that from the mixed method provides the best signal.

The lattice renormalization of the vector current can be defined from normalizing the vector charge,

$$g_{V_4}^b \equiv \frac{\text{Tr}[\Gamma^e \langle P | \int d^3x \bar{\psi}(x) \gamma_4 \psi(x) | P \rangle]}{\text{Tr}[\Gamma^e \langle P | P \rangle]} = \frac{1}{Z_V} \quad (36)$$

where superscript b is for bare value, and $\Gamma^e = (1 + \gamma_4)/2$ is the unpolarized projection operator. Fig. 6 shows that all the fitting methods mentioned in the last section provide consistent results, while the results from the “mixed” method have the best signals among the three methods. A constant fit for the cases with $m_\pi \in (0.25, 0.4)$ GeV gives the value of the vector renormalization factor as 1.096(6) which is just slightly smaller than the value 1.105(4) obtained from the axial Ward identity [41].

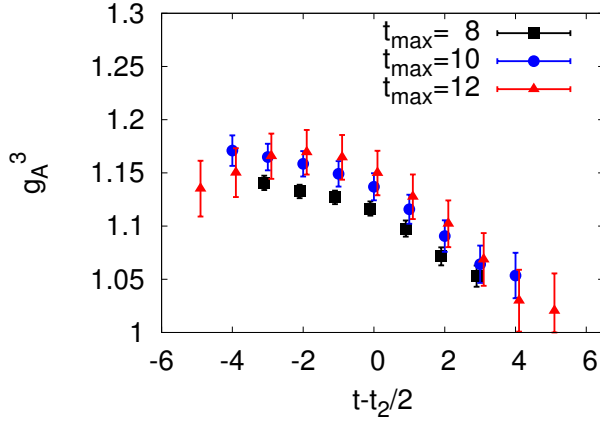


FIG. 7: The sink-source separation dependence of the matrix element of the isovector axial-vector current. The data of the isovector case with $t_2 = 8$ (black squares) are slightly smaller than that from the other two separations, while the result with $t_2=10$ (blue dots) is consistent with that with $t_2 = 12$ (red triangles).

Then the renormalization of the vector current can be used to renormalize the axial-vector matrix element with polarized projection,

$$\begin{aligned} g_A^b &\equiv \frac{\sum_{i=1,2,3} \text{Tr}[\Gamma_i^m \langle P | \int d^3x \bar{\psi}(x) \gamma_5 \gamma_i \psi(x) | P \rangle]}{3 \text{Tr}[\Gamma^e \langle P | P \rangle]} \\ g_A^R &\equiv g_A^b Z_V \\ &= \frac{\sum_{i=1,2,3} \text{Tr}[\Gamma_i^e \langle P | \int d^3x \bar{\psi}(x) \gamma_5 \gamma_i \psi(x) | P \rangle]}{3 \text{Tr}[\Gamma^e \langle P | \int d^3x \bar{\psi}(x) \gamma_4 \psi(x) | P \rangle]} \quad (37) \end{aligned}$$

where the superscript b/R stands for the bare/renormalized value respectively and $\Gamma_i^m = (1 + \gamma_4) \gamma_i \gamma_5 / 2$ is the polarized projection operator.

Using $g_{V_4}^b$ (instead of that from the axial Ward identity for pion) to renormalize g_A as in Eq. (37) could improve the signal of the renormalized g_A by $\sim 20\%$ since these two matrix elements are correlated. As observed

in Fig. 7, the sink-source separation dependence for the isovector case is mild, while a curve is observable at the right side of the plateau due to a larger excited state contribution from the point interpolation field at the sink. This is in contrast to the flatter behavior to the left of the plateau where the excited-state contribution is ameliorated by the smeared source. In Fig. 8, we plot the results of the isovector axial-vector coupling g_A^3 from the three fitting methods we mentioned. We note that those from the “mixed” method are always between those from the other two methods, for all the data points in the range of $m_\pi \in (0.25, 0.4)$ GeV. The values from the three methods at the unitary point are listed in Table II. Similar to other lattice calculations at this pion mass (i.e. ~ 300 MeV), irrespective of which fit is used, the isovector axial-vector matrix element, g_A^{u-d} is $\sim 10\%$ smaller than the experimental value 1.2723(23)[45].

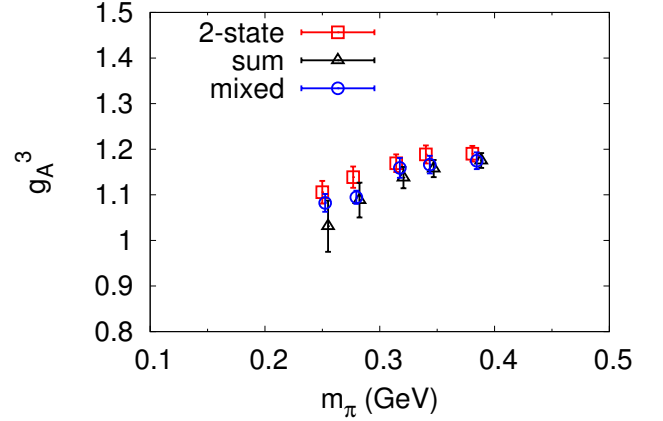


FIG. 8: The isovector axial-vector matrix element vs the pion mass, from three kinds of fitting method: 2-state fit (red squares), summed slope (black triangles), and the mixed fit which combines those two methods (the blue dots). The results from these different methods are consistent while that from the mixed method provides the best signal.

B. Scalar case

Similarly, the renormalized scalar matrix element with the unpolarized projection of the nucleon can be calculated by,

$$g_S \equiv \frac{Z_S \text{Tr}[\Gamma^e \langle P | \int d^3x \bar{\psi}(x) \psi(x) | P \rangle]}{\text{Tr}[\Gamma^e \langle P | P \rangle]}, \quad (38)$$

where the renormalization constant Z_S is obtained from the RI/MOM scheme and its value on the ensemble we use here is calculated to be 1.1397(54) [41]. On the other hand, if one just focuses on the $\pi N \sigma$ term, $2Z_m m_q Z_S g_S^b$, the renormalizations of the quark mass Z_m and that of

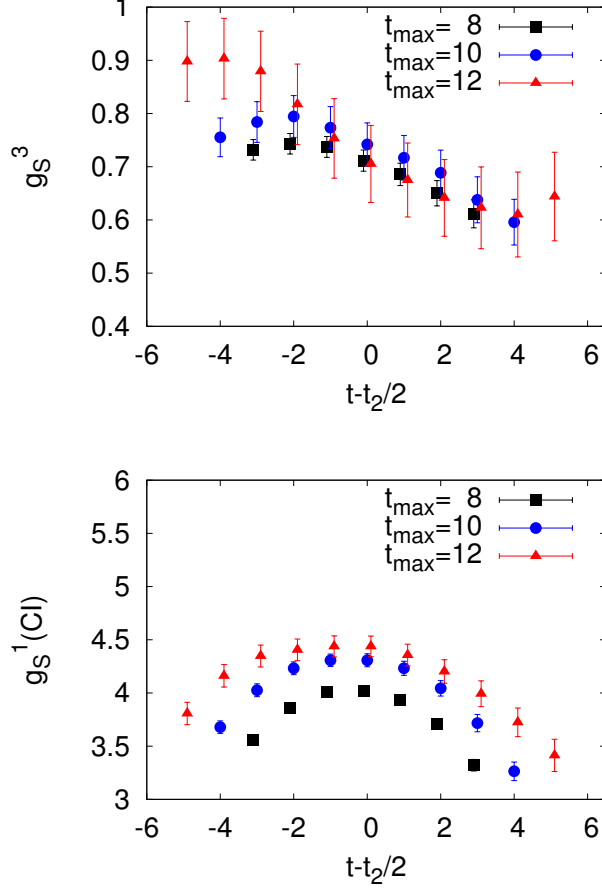


FIG. 9: The separation dependence of the matrix element of scalar current, for both the isovector and the CI part of the singlet case. The dependence is mild for the isovector case (the upper panel), while obvious for the CI part of the singlet case (the lower panel).

the scalar matrix element Z_S are canceled and so the $\pi N \sigma$ term is free of the renormalization.

It is interesting to point out that the CI part of the scalar singlet matrix element has a strong sink-source separation dependence, as seen in the lower panel of Fig. 9. At the same time, such a separation dependence seems to be canceled between the u and d quarks, so that the isovector case in the upper panel of Fig. 9 has only a mild separation dependence. The results for the isovector scalar matrix element from the three fitting methods are plotted in Fig. 10 and those at the unitary point are listed in Table II. This shows that, despite the fact that there are 2 u valence quarks and only one d quark in the proton, the d contribution to the scalar matrix element per quark is more than that of the u , as

$$\frac{g_{S,CI}^u}{2g_{S,CI}^d} = 0.67(2) \quad (39)$$

is much smaller than one. The scalar matrix elements of

both the u and d quark increase as m_q decreases, but the isovector scalar matrix element is not far from unity over the entire quark mass region from light to heavy. This has been interpreted to be related to the Gottfried sum rule violation [46] where it is found experimentally that there are more d antipartons than u antipartons..

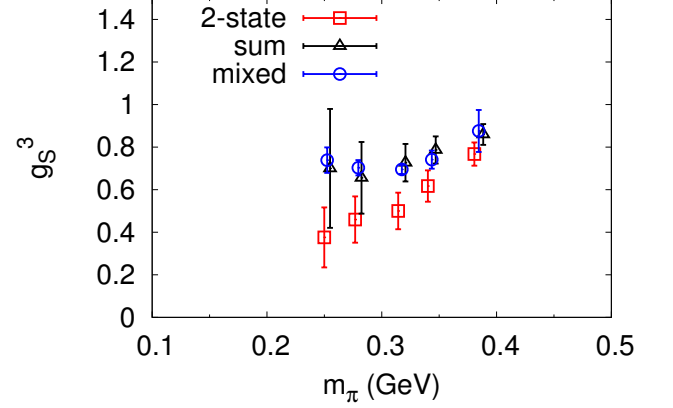


FIG. 10: Isovector scalar matrix element vs. pion mass from three kinds of fitting method: 2-state fit (red squares), summed slope (black triangles), and the mixed fit which combines those two methods (blue dots). The results from these different methods are slightly different.

C. Quark momentum fraction

The quark momentum fraction in the nucleon can be calculated with the traceless part of the energy momentum tensor, and it should be consistent between calculations with two different operators. The first one uses the combination of the diagonal temporal and spatial components of the energy momentum tensor,

$$\langle x \rangle^E \equiv \frac{\text{Tr}[\Gamma^e \langle P | \int d^3x O^E(x) | P \rangle]}{E \text{Tr}[\Gamma^e \langle P | P \rangle]}. \quad (40)$$

where $O^E(x) = \bar{\psi}(x) \frac{1}{2} (\gamma_4 \overleftrightarrow{D}_4 - \frac{1}{3} \sum_{i=1,2,3} \gamma_i \overleftrightarrow{D}_i) \psi(x)$ is the traceless part of the energy momentum tensor T_{44} and is a measure of the quark fraction of the nucleon mass or energy. The related matrix element can be calculated in the rest frame and, as a result, it will have a good signal. On the other hand, the operator T_{44} itself can have mixing with lower dimension operators like the dimension-3 scalar operator $\bar{\psi}(x)\psi(x)$. Nevertheless, such a mixing will be canceled due the subtraction of the diagonal spatial components in O_E .

The other approach uses the forward off-diagonal matrix components of the energy momentum tensor (T_{4i}) in

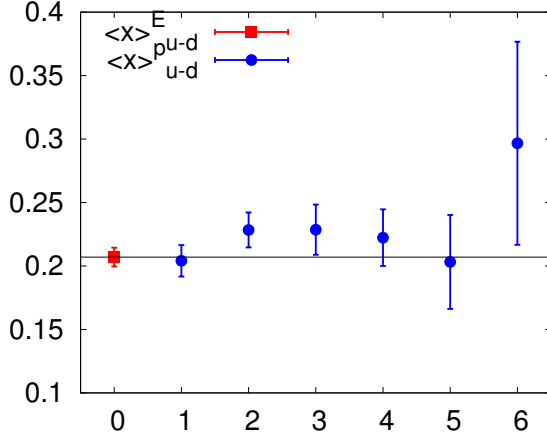


FIG. 11: The plateau fit values ($t_2 = 10$ case) of the isovector momentum fraction for $\langle x \rangle_E$ in the rest frame (red square) and also $\langle x \rangle_P$ in a moving frame with different momenta (blue dots). The results from both the diagonal and off-diagonal components (and also that from different momenta based on the off-diagonal matrix components) are consistent, but the first approach provides much better SNR.

a moving frame,

$$\langle x \rangle^P \equiv \frac{\text{Tr}[\Gamma^e \langle P | \int d^3x \bar{\psi}(x) \frac{1}{4} (\gamma_i \overleftrightarrow{D}_4 + \gamma_4 \overleftrightarrow{D}_i) \psi(x) | P \rangle]}{p_i \text{Tr}[\Gamma^e \langle P | P \rangle]} \quad (41)$$

with p_i being the i -th component of the nucleon momentum. Therefore, it is a measure of the quark momentum fraction in a moving nucleon. Such a scheme is free of mixing of the lower dimension operators due to its tensor structure, while the corresponding matrix element is proportional to the momentum and is thus more noisy than that from the first approach, because mixed momentum sources are involved for the matrix element of the nucleon at non-zero momentum.

Fig. 11 shows the plateau fit values of the $t_2=10$ case for the quark isovector momentum fraction. They are $\langle x \rangle_E$ from the diagonal components of the energy-momentum tensor with the nucleon in the rest frame and also $\langle x \rangle_P$ from the off-diagonal components in a moving frame with different momenta. The results from both the diagonal and off-diagonal components (and also those from different momenta) are consistent, but $\langle x \rangle_E$ provides much better SNR. The sink-source separation dependence is shown in Fig. 12, for both results based on the diagonal components and off-diagonal components. It is interesting to observe that the separation dependence of the isovector quark momentum fraction based on the off-diagonal components seems to be milder than that based on the diagonal ones, for the cases with $t_2=8$ and 10. The $\langle x \rangle_P$ case with $t_2=12$ seems to have some t dependence at the smeared source end, but it could be due to the statistical fluctuation due to relatively poor

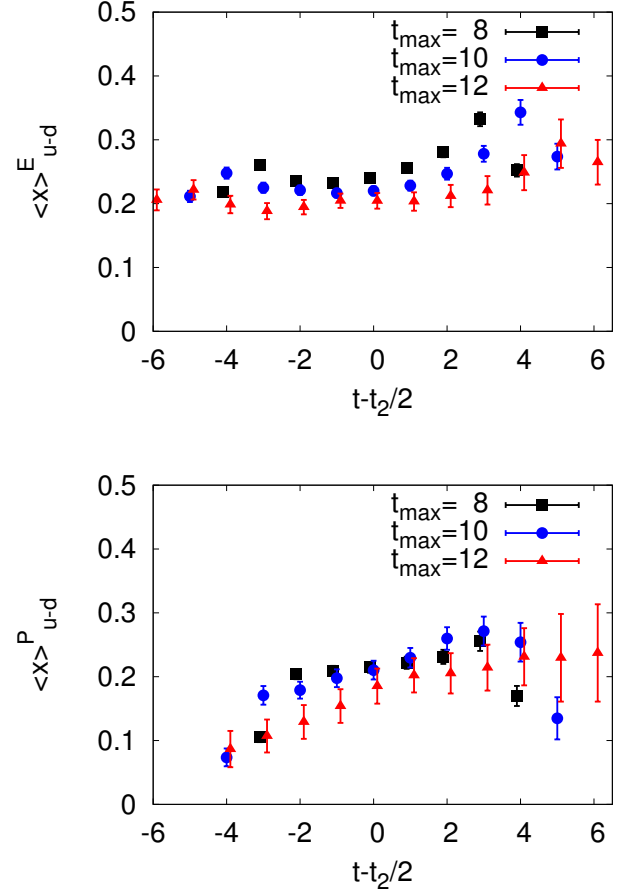


FIG. 12: The sink-source separation dependence of the isovector quark momentum fraction for the case of the diagonal components of the energy momentum tensor (the upper panel) and that of the off-diagonal ones (the lower panel).

signal.

As in Ref. [51], the renormalization factor for the ensemble we used has been obtained with the one-loop lattice perturbative theory, as 1.049(3), in the \overline{MS} scheme at 2 GeV. The error is from the uncertainty of the lattice spacing. The renormalized values of the isovector quark momentum fraction of $\langle x \rangle_E$ from the three fitting methods are plotted in Fig. 13, and those at the unitary point are listed in Table II.

VI. SUMMARY

We have introduced a new method to calculate the nucleon matrix elements in the connected insertion. The stochastic sandwich method (SSM) with low-mode substitution (LMS) is an approach which uses low modes for the all-to-all quark propagator between the current and the sink and the corresponding high-mode contribution is taken care of by the noise propagator from the sink

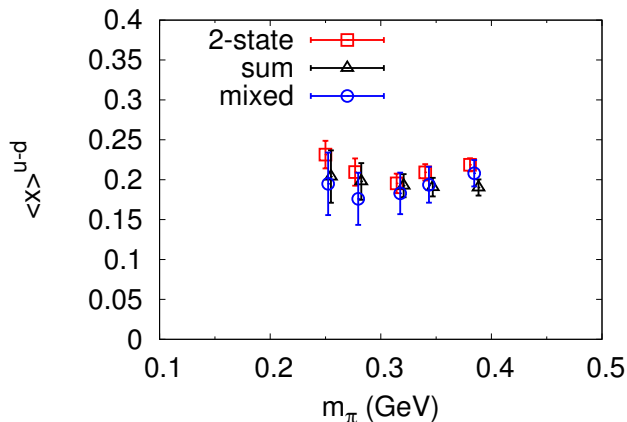


FIG. 13: The isovector quark momentum fraction ($\langle x \rangle^E$ in the rest frame) vs. the pion mass, from three fitting methods: 2-state fit (red squares), summed slope (black triangles), and the mixed fit which combines those two methods (blue dots). The results from the three methods are consistent.

to the current. We have shown that it is more efficient than the sink- and current- sequential methods. However, it does not scale well with volume which requires more low eigenmodes. It will lose its advantage when the overhead from calculating the LMS for all the quark propagators involved is more than the amount it saves compared with the sink-sequential or current-sequential method. But this will occur only at volumes much larger than that used here.

We have used three fitting methods. One is a two-state fitting including the contamination from the excited-state transition and the second is the summed-slope method. The third is a mix of these two methods.

The proton isovector axial-vector coupling g_A^3 we obtain with the overlap fermion at the unitary point with $m_\pi=330$ MeV is

$$g_A^3 = 1.166(19) \quad (42)$$

which is $\sim 8\%$ smaller than the experimental value.

The separation dependence of this quantity is mild. Since it is smaller than the experimental value on this lattice, it is essential to repeat the calculation of g_A^3 on larger volumes and with lighter quark masses.

For the isovector scalar matrix element in the proton, the renormalized value at $\overline{MS}(2\text{GeV})$ at the unitary point is

$$g_S^3 = 0.74(4). \quad (43)$$

This shows that, despite the fact that there are 2 u valence quarks and only one d quark in the proton, the d contribution to the scalar matrix element per quark is more than that of the u , as

$$\frac{g_{S,CI}^u}{2g_{S,CI}^d} = 0.67(2) \quad (44)$$

is much smaller than one. This has been interpreted [46] to be related to the Gottfried sum rule violation [47] where it is found experimentally that there are more d antipartons than u antipartons.

In the isovector quark momentum fraction case, the bare value we obtained at the unitary point on the ensemble mentioned above is

$$\langle x \rangle_{u-d} = 0.192(19), \quad (45)$$

with the renormalization factor 1.049(3) from one-loop lattice perturbative theory [51]. This value is similar to those from most lattice calculations [3, 7, 14–16] and is larger than the experimental value. However, the $O(a^2)$ error has not been considered. It can be assessed by imposing the momentum and angular momentum sum rules at finite lattice spacing as is demonstrated in a quenched calculation [52]. We will return to this issue when the complete lattice simulation of the momentum and angular-momentum decompositions is carried out.

We will perform calculations with physical sea quark masses in the future.

Acknowledgments

We thank the RBC and UKQCD Collaborations for providing us their DWF gauge configurations. This work is supported in part by the U.S. Department of Energy under Grant No. DE-FG05-84ER40154, and desc0013065. A.A. acknowledges the support of NSF CAREER through grant PHY-1151648. M.G. is partially supported by the National Science Foundation of China (NSFC) under the project No. 11405178 and the Youth Innovation Promotion Association of CAS (2015013). This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

[1] B. J. Owen, J. Dragos, W. Kamleh, D. B. Leinweber, M. S. Mahbub, B. J. Menadue and J. M. Zanotti, Phys.

Lett. B **723**, 217 (2013) [arXiv:1212.4668 [hep-lat]].
[2] T. Bhattacharya, S. D. Cohen, R. Gupta, A. Joseph,

- H. W. Lin and B. Yoon, Phys. Rev. D **89**, no. 9, 094502 (2014) [arXiv:1306.5435 [hep-lat]].
- [3] C. Alexandrou, M. Constantinou, S. Dinter, V. Drach, K. Jansen, C. Kallidonis and G. Koutsou, Phys. Rev. D **88**, 014509 (2013) [arXiv:1303.5979 [hep-lat]].
- [4] C. Alexandrou *et al.* [ETM Collaboration], Phys. Rev. D **83**, 045010 (2011) [arXiv:1012.0857 [hep-lat]].
- [5] S. Ohta [RBC and UKQCD Collaborations], PoS LATTICE **2013**, 274 (2014) [arXiv:1309.7942 [hep-lat]].
- [6] J. D. Bratt *et al.* [LHPC Collaboration], Phys. Rev. D **82**, 094502 (2010) [arXiv:1001.3620 [hep-lat]].
- [7] S. Syritsyn *et al.*, PoS LATTICE **2014**, 134 (2015) [arXiv:1412.3175 [hep-lat]].
- [8] J. R. Green, M. Engelhardt, S. Krieg, J. W. Negele, A. V. Pochinsky and S. N. Syritsyn, Phys. Lett. B **734**, 290 (2014) [arXiv:1209.1687 [hep-lat]].
- [9] S. Capitani, M. Della Morte, G. von Hippel, B. Jager, A. Jüttner, B. Knippschild, H. B. Meyer and H. Wittig, Phys. Rev. D **86**, 074502 (2012) [arXiv:1205.0180 [hep-lat]].
- [10] R. Horsley, Y. Nakamura, A. Nobile, P. E. L. Rakow, G. Schierholz and J. M. Zanotti, Phys. Lett. B **732**, 41 (2014) [arXiv:1302.2233 [hep-lat]].
- [11] G. S. Bali *et al.*, Phys. Rev. D **91**, no. 5, 054501 (2015) [arXiv:1412.7336 [hep-lat]].
- [12] A. Abdel-Rehim *et al.*, arXiv:1507.04936 [hep-lat].
- [13] Y. Aoki, T. Blum, H. W. Lin, S. Ohta, S. Sasaki, R. Tweedie, J. Zanotti and T. Yamazaki, Phys. Rev. D **82**, 014501 (2010) [arXiv:1003.3387 [hep-lat]].
- [14] G. S. Bali *et al.*, Phys. Rev. D **90**, no. 7, 074510 (2014) [arXiv:1408.6850 [hep-lat]].
- [15] D. Pleiter *et al.* [QCDSF/UKQCD Collaboration], PoS LATTICE **2010**, 153 (2010) [arXiv:1101.2326 [hep-lat]].
- [16] Y. Aoki, T. Blum, H. W. Lin, S. Ohta, S. Sasaki, R. Tweedie, J. Zanotti and T. Yamazaki, Phys. Rev. D **82**, 014501 (2010) [arXiv:1003.3387 [hep-lat]].
- [17] J. R. Green, J. W. Negele, A. V. Pochinsky, S. N. Syritsyn, M. Engelhardt and S. Krieg, Phys. Rev. D **86**, 114509 (2012) [arXiv:1206.4527 [hep-lat]].
- [18] M. Gonzalez-Alonso and J. Martin Camalich, Phys. Rev. Lett. **112**, no. 4, 042501 (2014) [arXiv:1309.4434 [hep-ph]].
- [19] T. Bhattacharya, V. Cirigliano, S. D. Cohen, A. Filipuzzi, M. Gonzalez-Alonso, M. L. Graesser, R. Gupta and H. W. Lin, Phys. Rev. D **85**, 054512 (2012) [arXiv:1110.6448 [hep-ph]].
- [20] Y. Aoki *et al.* [RBC and UKQCD Collaborations], Phys. Rev. D **83**, 074508 (2011) [arXiv:1011.0892 [hep-lat]].
- [21] A. Li *et al.* [xQCD Collaboration], Phys. Rev. D **82**, 114501 (2010) [arXiv:1005.5424 [hep-lat]].
- [22] A. Alexandru, M. Lujan, C. Pelissier, B. Gamari and F. X. Lee, arXiv:1106.4964 [hep-lat].
- [23] T. A. DeGrand and S. Schaefer, Comput. Phys. Commun. **159**, 185 (2004) [hep-lat/0401011].
- [24] L. Giusti, P. Hernandez, M. Laine, P. Weisz and H. Wittig, JHEP **0404**, 013 (2004) [hep-lat/0402002].
- [25] L. Giusti, P. Hernandez, M. Laine, C. Pena, J. Wenekers and H. Wittig, Phys. Rev. Lett. **98**, 082003 (2007) [hep-ph/0607220].
- [26] J. Foley, K. Jimmy Juge, A. O'Cais, M. Peardon, S. M. Ryan and J. I. Skullerud, Comput. Phys. Commun. **172**, 145 (2005) [hep-lat/0505023].
- [27] T. Kaneko *et al.* [JLQCD Collaboration], PoS LAT **2007**, 148 (2007) [arXiv:0710.2390 [hep-lat]].
- [28] M. Gong [XQCD Collaboration], A. Alexandru, Y. Chen, T. Doi, S.J. Dong, T. Draper, W. Freeman, M. Glatzmaier, A. Li, K.F. Liu, and Z. Liu, Phys. Rev. D **88**, no. 1, 014503 (2013) [arXiv:1304.1194 [hep-ph]].
- [29] W. Wilcox, T. Draper and K. F. Liu, Phys. Rev. D **46**, 1109 (1992) [hep-lat/9205015].
- [30] T.-W. Chiu, Phys. Rev. D **60**, 034503 (1999) [hep-lat/9810052].
- [31] K.-F. Liu and S.J. Dong, Int. J. Mod. Phys. A **20**, 7241 (2005) [hep-lat/0206002].
- [32] T.-W. Chiu and S. V. Zenkin, Phys. Rev. D **59**, 074501 (1999) [hep-lat/9806019].
- [33] C.W. Bernard, Gauge Theory on a Lattice, 1984, edited by C. Zachos *et al.*, Argonne National Laboratory, Argonne, IL (1984) 85; T. Draper, Ph. D. thesis, UMI-84-28507 (1984); C. W. Bernard, T. Draper, G. Hockney, A. M. Rushton and A. Soni, Phys. Rev. Lett. **55**, 2770 (1985).
- [34] G. Martinelli and C. T. Sachrajda, Nucl. Phys. B **316**, 355 (1989).
- [35] T. Draper, R. M. Woloshyn and K. F. Liu, Phys. Lett. B **234**, 121 (1990).
- [36] T. Draper, R. M. Woloshyn, W. Wilcox and K. F. Liu, Nucl. Phys. B **318**, 319 (1989).
- [37] R. Evans, G. Bali and S. Collins, Phys. Rev. D **82**, 094501 (2010) [arXiv:1008.3293 [hep-lat]].
- [38] G. S. Bali *et al.*, PoS LATTICE **2013**, 271 (2014) [arXiv:1311.1718 [hep-lat]].
- [39] C. Alexandrou *et al.* [ETM Collaboration], Eur. Phys. J. C **74**, no. 1, 2692 (2014) [arXiv:1302.2608 [hep-lat]].
- [40] T. Blum *et al.* [RBC and UKQCD Collaborations], arXiv:1411.7017 [hep-lat].
- [41] Z. Liu *et al.* [chiQCD Collaboration], Phys. Rev. D **90**, no. 3, 034505 (2014) [arXiv:1312.7628 [hep-lat]].
- [42] Y. B. Yang *et al.*, arXiv:1410.3343 [hep-lat].
- [43] L. Maiani, G. Martinelli, M. L. Paciello and B. Taglienti, Nucl. Phys. B **293**, 420 (1987).
- [44] M. Deka, T. Streuer, T. Doi, S. J. Dong, T. Draper, K. F. Liu, N. Mathur and A. W. Thomas, Phys. Rev. D **79**, 094502 (2009) [arXiv:0811.1779 [hep-ph]].
- [45] K. A. Olive *et al.* [Particle Data Group Collaboration], Chin. Phys. C **38**, 090001 (2014).
- [46] K. F. Liu and S. J. Dong, Phys. Rev. Lett. **72**, 1790 (1994) [hep-ph/9306299].
- [47] P. Amaudruz *et al.* [New Muon Collaboration], Phys. Rev. Lett. **66**, 2712 (1991).
- [48] Y. B. Yang, M. Gong, K. F. Liu and M. Sun, PoS LATTICE **2014**, 138 (2014) [arXiv:1504.04052 [hep-ph]].
- [49] P. Hasenfratz, S. Hauswirth, T. Jorg, F. Niedermayer and K. Holland, Nucl. Phys. B **643**, 280 (2002) [hep-lat/0205010].
- [50] P. A. Boyle, arXiv:1411.5728 [hep-lat].
- [51] M. Glatzmaier, in preparation.
- [52] M. Deka *et al.*, Phys. Rev. D **91**, no. 1, 014505 (2015) [arXiv:1312.4816 [hep-lat]].